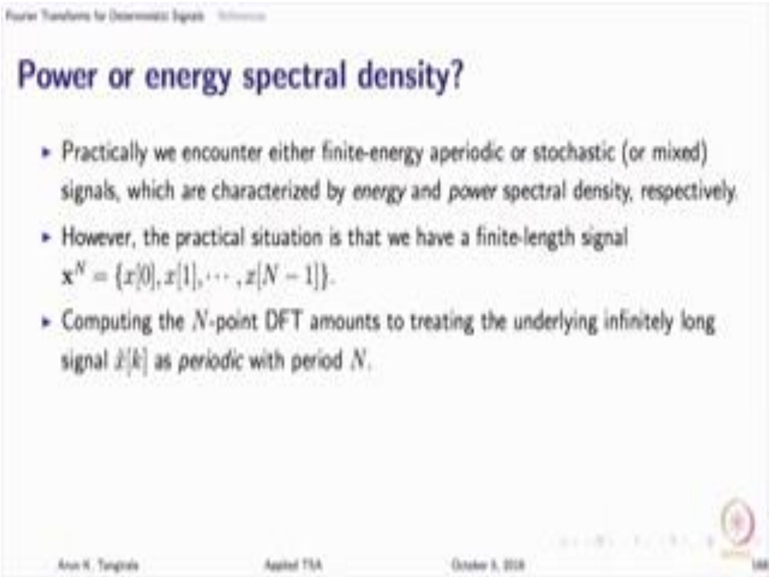
Applied Time-Series Analysis
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture - 70
Lecture 31B - DFT and Periodogram 3 (with R Demonstrations)

(Refer Slide Time: 00:11)

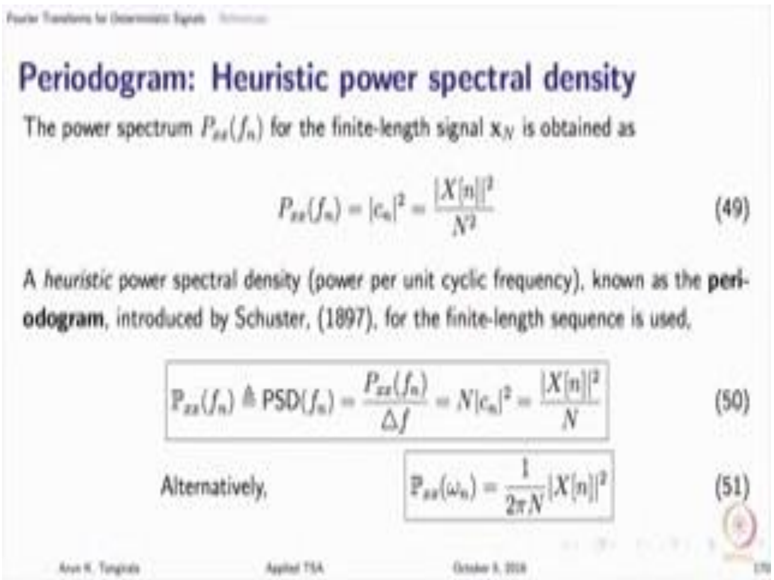


And now we come to the point of power I mean the periodogram, and before we jump into that we should ask whether it is appropriate to talk of power spectral density or power spectrum or energy spectral density, what is the appropriate term to use because I do not know what the underlying signal is, right whatever I observed it could be observations from some exponential decaying signal for example, or some wearied signal which is not periodic. Then I should be constructing energy spectral density if it is of finite energy may be it is periodic, then I should be constructing power spectrum, but I do not know and DFT says I also do not know, but what I am going to force is, I am going to force the periodicity nature on it, I do not care what the underlying signal is, I am going to assume it to be periodic therefore, I will always give power spectrum to you, I will always imagine the signal to be a periodic.

So, when you plot this squared DFT coefficients, you would be plotting the power spectrum, but very often you are interested in densities also. So, what do you do? can you define power spectral density for a periodic signal no therefore, what one does is constructs an empirical power spectral density this is empirical because the truth we do not know what it is, why do we want to construct power spectral, this empirical power spectral density? Because what is the underlying signal is not periodic may be spectral densities can make sense.

If it is periodic then it is an empirical one. So, it may be worthwhile always working with this pi empirical power spectral density and it becomes very useful also when we move to the random signal world because we mo we know random signals are power signals and hopefully they have a power spectral density. And if they do at least periodogram will give me some estimate of the power spectral density, it is with all of this motivation and some futuristic vision, that Custer in really you know late 1890s on the verge of the don of 20th century.

(Refer Slide Time: 02:40)



That he introduced this concept known as periodogram, this is nearly 100 years after Fourier had come up with this Fourier transform and I am sure you know Fourier would have never ever thought that the box that is opening has so many things to offer and that it will haunt millions of people, millions of students let us say because of it is time domain I understand everything, frequency domain you know I. So, that is what you frequently hear, ok anyway.

So, this periodogram is what is at the heart of all spectral analysis tools and now you should have a clear picture of what is periodogram. Periodogram is an empirical spectral density, what is meant by spectral density? Power per unit frequency correct; so first understand what is power? We are assuming DFT assumes the signal to be periodic correct? It assumes the signal to be periodic therefore, what I would do is I would compute the discrete time Fourier series coefficient of course, there is a relation between that and DFT we have already seen that.

(Refer Slide Time: 03:55)

So, assuming that the signal is periodic with length and which is what DFT assumes, I would compute C n the discrete time Fourier coefficients and then how does one construct the power spectrum from Parseval's relation, we already know from the discussion on DTFs we know that the power spectrum of this periodic signal of length n is given by mode C n square. We know we know that correct and by virtue of the relation between C n and the DFT coefficient what is the relation between C n and X of n what? C n is.

Student: x n.

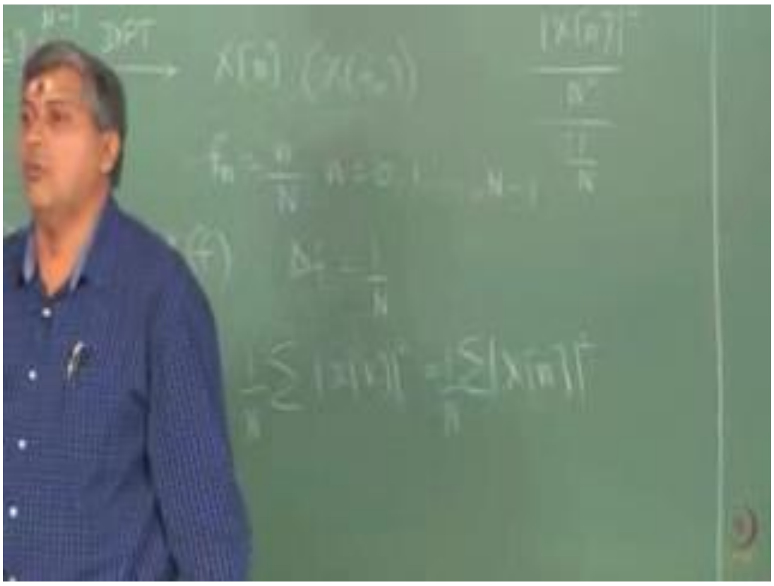
1 over n x n correct and that is what we have used to strike the relation in 49. Now periodogram is this power per unit frequency, it is empirical one; how many frequencies are we computing the DFT over?

Student: n frequencies.

N frequencies no n may run from n runs from 0 to n minus 1, but in total you have n frequency points. So, you say or you can look at this way you can say that it is psd per delta f right and delta what is delta f in DFT? 1over n, so the periodogram is mod x of n square by n square this is the power per delta f right per unit change in the frequency delta f and delta f is 1 over N as a result your periodogram is 1 over N square DFT coefficient that is what you can think of and that is what is Custer definition is and people have been we have been using this definition now for nearly 115 year.

You this is when you write in the sense now you have to be careful.

(Refer Slide Time: 06:00)

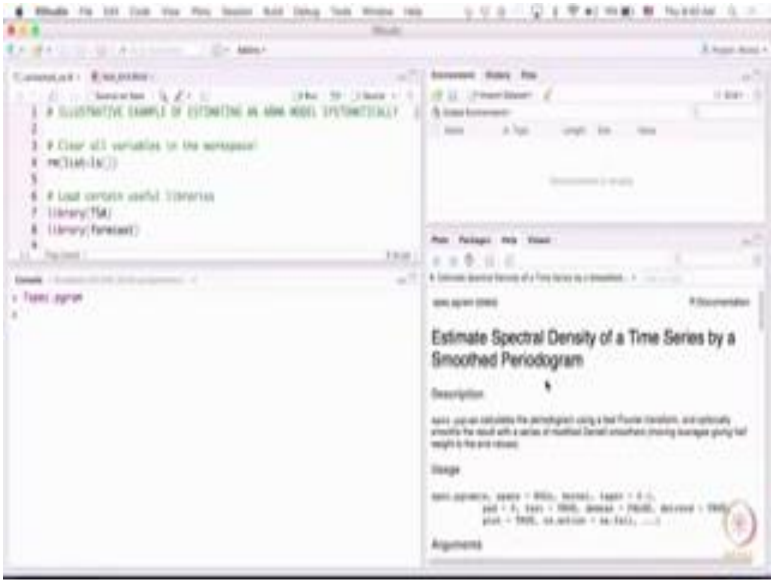


Because now this is power per unit frequency, you have to be careful with whether you are working with cyclic frequencies or angular frequencies right when I talk of densities let us say volume when I am talking of densities, it does matter whether I am talking of kg per cubic meter or kg per cc or grams per cc and so on. Here the numerator units do not change when you go from cyclic to angular, the denominator that is delta f changes therefore, you have to be careful whether the periodogram which is the empirical power spectral density is being expressed as power per cyclic frequency or power per angular frequency.

If you choose to work with the angular frequencies then you need to have a 2 pi factor there that is pretty straight forward. Again you should be careful it is your responsibility when you use a periodogram routine or something that claims to compute power spectral density like in matlab it says psd, I pretty all the users who just blindly use psd without knowing the basic fact that it is not the true psd, it is just an empirical psd and the underlying math is what we have seen here, you have to be careful and you have to go and check whether the psd is being reported in terms of angular frequency or cyclic frequency.

In matlab it is reported in terms of angular frequency whereas, r for example, has this routine periodogram in TSA.

(Refer Slide Time: 07:35)

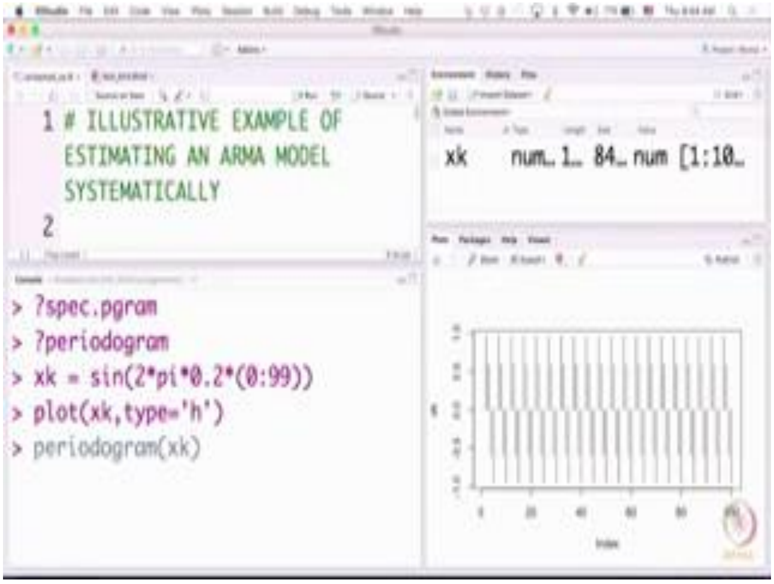


You also have, I am going to start from a clean slate fine. So, r has 2 routines for use spec dot pgram which estimates spectral density of a time series by smooth periodogram you have to be careful we have not yet gone there, the smooth periodogram concept we will discuss when we get into estimation, at the moment we have not yet gotten the estimation.

So, this spec dot pgram as a number I do not know how many can actually see, but first let me get to the jumbo font all of you are able to see at least the command window font. I do not know how are you are able to see the help font here, but you see that of course, it asks for the signal and then there are a number of other options. Although options pertain to smoothing that would be necessary when you use this for a random signal for right now we are only talking of the deterministic signals.

So, I would prefer to use the periodogram from the TSA, this is very simple, it computes exactly what we have shown on the slide; if there is no smoothing or anything done here this is called a raw periodogram in the estimation domain.

(Refer Slide Time: 09:01)



So, it just asks for unit time series and gives and whether you want to plot the logarithm of the power which is essentially it turns out to be decibels, 20 log 10 of power is decibels and you know decibel is named after Alexander graham bell.

So, lot times the logarithm of power as a better relationship and insight to show. So, many a times people plot decibels rather than power right, you also say the decibel level is low and so on, the other day I had ordered a Bluetooth speaker from Philips and as just going through the specifications and it had all the things that we have learned the bandwidth the frequency range the you know the decibel value what is everything. So, it would be now probably you would be able to you dare to read the manual of any communications device or any this kind of speakers and so on, you will understand better at least.

So, you have the option of plotting either the decibel or the power, the default option is just the raw one absolute value and I like it that way do not give me decibels and so on because if necessary we will do the logarithmic. So, for example, if you were to do the sin take the just generate a sin wave right, let us pick a frequency, we cannot pick any frequency we have to pick frequencies in an interval minus half to half or 0 to 1 and so on. So, let us pick 0.2 right and generate let us say about 100 samples right, alright. So, what is the period of this signal? Frequency is 0.2 what is the period?

Student: (Refer Time: 10:45).

10 is correct and should I have asked what is the fundamental period? If you are a lawyer you would say yeah 10 is also correct, but illegally 10 is correct, but let me a fundamental period sometimes fundamental can really go make you go mental fine. So, this is what your sinusoid is; better to always plot and see if it is a same signal that I have generated or something else. Let us generate a kind of bar or stem plot it looks like a sine wave you can of course, zoom in and see we Would not go into that let us quickly get to the periodogram right what would you expect to see?

Student: peaks are fundamental (Refer Time: 11:28).

Peaks all over.

Student: and fundamental (Refer Time: 11:34).

What is that fundamental? So, first of all what is periodogram going to compute it is going compute your DFT as you see on the board am I right? How many coefficients is it going to calculate, what is the length of the signal that we have here?

Student: 100.

100. So, it is going to compute 100 DFT coefficients and what do you expect the periodogram to show as it is going to give me mod X of n square over N correct.

(Refer Slide Time: 12:03)

I hope so does you can go through the help here and see if it is indeed giving you that. So, if you see here I mean all though it does not say exactly mod of x of n square over N, but it also by the way it gives you the frequencies all of it is returned. So, let us actually return this into some variable yes I can say x k per. So, it reports I mean some breathing spaces could fine. So, what is it going to report? It is going to report the periodogram and the frequencies at which it is computing. When it is going to plot, it is going to plot only for one half because it is symmetric, what do we expect to see now? So, it is going to compute at frequencies 0 and then 1 over 100, 2 over 100, 3 over 100 up to 999 over 100 what do we expect to see in the periodogram?

Student: (Refer Time: 13:07).

Why?

Student: Peak.

Peak at?

Student: (Refer Time: 13:15).

What is the corresponding value of small n?

Student: 20.

Somebody says 20; somebody says 19 and remembers r starts counting form 1, theory counts from 0, right? So, which coefficient do we expect to see a peak? In fact, what do you should do is as a simple home work although I am going to show you periodogram, take the same sine wave and do an FFT in r the algorithm is FFT, when you do an FFT you would get the coefficients and you should check what the which coefficient do we expect to see to be non zero and which are the ones that you expect to be 0. So, which coefficient do you expect to be to see as non zero here for this example?

Student: 20.

20.

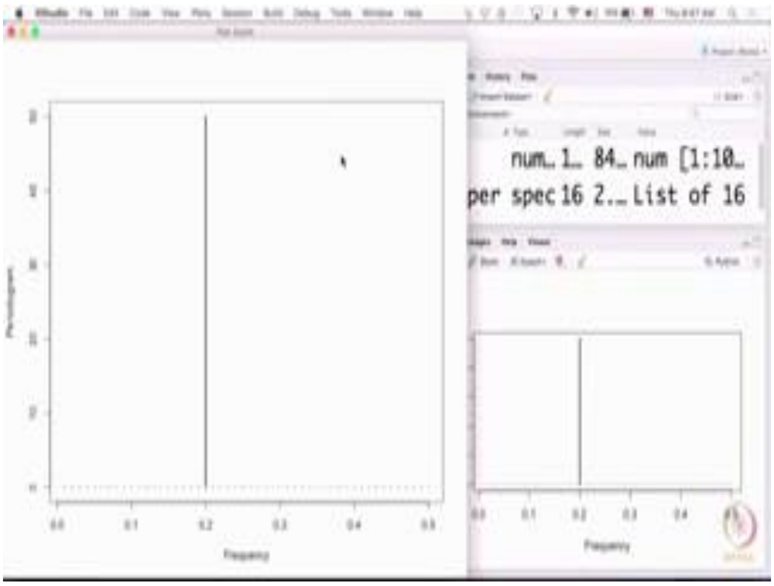
Student: (Refer Time: 14:10).

Sure, but in r language.

Student: 20 (Refer Time: 14:15).

21 anyway for we will come back to that qualitatively what do we expect to see a peak in the periodogram at just a single n and the 0 everywhere correct very good. So, let us see we are right at that is it you are passed the Fourier test correct.

(Refer Slide Time: 14:43)



Do you like this? So, without instead of breaking your head on looking at this sine wave visually, just look at the periodogram I will tell you what, whether there is periodicity in it and if it is what are the frequencies that are present that is the beauty right. So, you can think of Fourier always this example is given, you can think of Fourier transform like a prism like it what does what does that prism does do? It actually takes white light and breaks it up into different colors do you recall yes and often colors are characterized by frequency bands right that is exactly. So, there also you cannot escape frequency, white light is supposed to contain all frequencies; I am giving you some hints on white noise alright, so now.

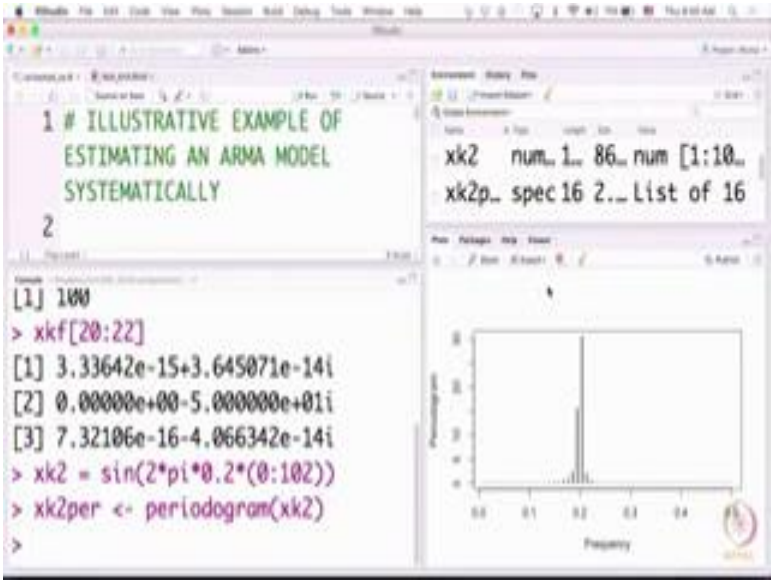
Student: Sir (Refer Time: 15:36).

Why?

Student: To conjugate (Refer Time: 15:40).

No, no we are only plot plotting the one half; if you were to plot the full periodogram there will be another peak you are right. So, you should go back and do periodogram is doing all of that for you, but it is a good idea for you to do it manually compute FFT, so the if you were to do that you can actually compute FFT of x k right x k f it contains the FFT and what it be expect to see the 21st coefficient first of all.

(Refer Slide Time: 16:12)



What is a length of x k f 100 correct? So, it is computed and we expect the 21st coefficients to be the maximum and the rest of them numerically very very small.

So, let us ask the coefficients what do you see does it confirm or expectation? What do you see there is the DFT coefficients this n and the n in r differ only by 1. So, we expect the 21st coefficient to be incidentally we are also in the 21st century. So, twenty first coefficients to be high and the 20th and the 22nd ones are extremely small they are just numerically very small do you see that.

Now, let us actually repeat this for another the same signal, but just change the length of the signal we are going to let us do that; any questions until now fine. So, to your question you do your at x k f and plot, by the way if you were to plot periodogram by yourself you have to create the f factor and you know exactly how to create. So, plot and then you will see two peaks that is the good idea to do it all right. So, x k this is how we have generated, what we are going to do is we are going to say that I have 3 more observations with me and call it x k 2 and also now recomputed the periodogram, the same signal I have more observations now it seems to have some siblings and cousins this peak right does it make sense? So, same sine wave have you have you change the frequency, have I added more frequencies I have not right. So, why should the periodogram tell me that look there is not only did it exactly hit point 2? It does not. On top of it, it is says there are all these other frequencies that are present they are not tiny ones how do you explain this?

Student :(Refer Time: 18:38).

What discontinuity did I have been?

Student: So, by adding 3 additional observation (Refer Time: 18:47) some ten periods (Refer Time: 18:48).

Ok.

Student: Initially you had some n periods.

Integer number of periods in the length of the observation.

Student: Now you have (Refer Time: 18:58) number and then and if it assumes periodic (Refer Time: 19:00).

Ok.

Student: Then there will be a discontinuity at that point.

You should not be sorry you actually helped him.

Student: Should have been a discontinuity and he made all kind of (Refer Time: 19:09) frequencies to (Refer Time: 19:13).

Good you are going to say something.

Student: (Refer Time: 19:18).

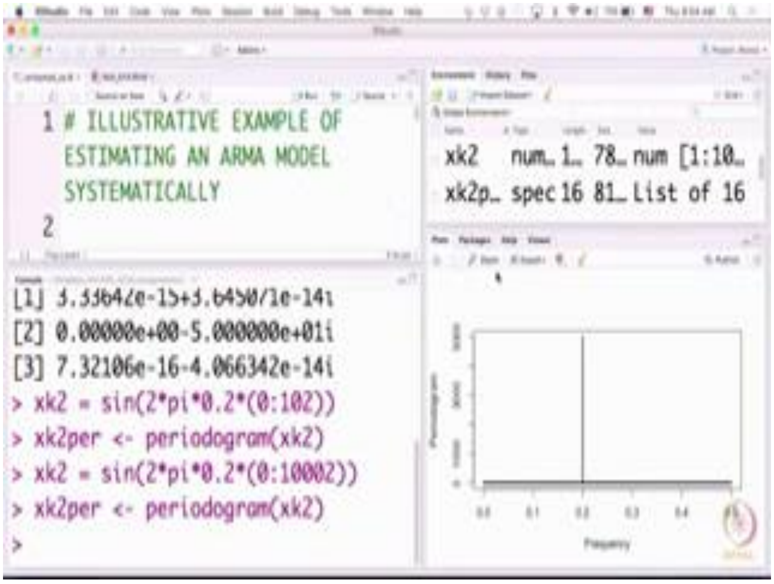
Yeah that does not, no no the grid is final as got, I can a go from zero to nine I mean 100 samples to 1000. I will still see only one peak; the fact is this signal as he just explained as complete and not completed the full cycles that is one perspective right it has completed fractional cycles, but your DFT assumes that it has completed integer number of cycles that is what periodicity assumption means. It is assuming that your signal is periodic outside the observation interval and that was perfectly valid when we had 100 observations, that is why I asked you what is a period; a period is 5. So, it completed 20 cycles earlier, now it has not completed integer number of cycles, it is 20 plus a fraction and then DFT is assuming it is periodic with that it isn't.

So, in the earlier case whatever periodicity DFT assumed was perfectly valid right it assume DFT assume that the period was your 100 and which was fine it is not a fundamental period, but that is, but now DFT assumes the period is 102 or 103 rather right that is not the case, which means in your basis bank you have building blocks that do not have the same period none of them as the original signal. Therefore, there is a mismatch and this is always going to be more or less going to be the case, how can you guaranty that you have collected data such that there are integer number of cycles, if that were the case you would know the period already.

Here we are saying in reality I do not know whether the signal is periodic and even if it is I do not know what the period is. DFT is helping me answer both these questions. So, what we see here is so called spectral leakage. Ideally I should have expected to see a peak at point 2, but 0.2 is not being hit at all here, in the grid point the point is not being hit. So, that is why the neighboring sinusoids are coming to the rescue and trying to explain this point. The other way of looking at it is the power at 0.2 now has leaked to neighboring frequencies and that is why it is called spectral leakage.

Typically there are two ways in which you can mitigate spectral leakage: one is to collect more and more samples, so that the fractional component does not have so much of an impact, for example here if very quickly just a few seconds I change.

(Refer Slide Time: 22:00)



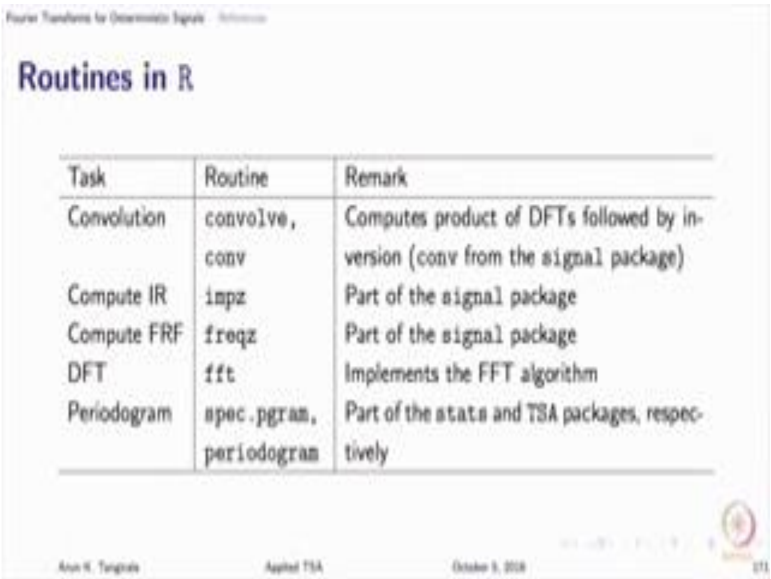
This x k 2 to be of not length 102, but 10002; even now it is fractional, but the number of integers cycles are a lot more and now if I do this do you see that the spectral leakage has vanished although it has not strictly hit 0.2, but this spectral leakage has gone.

But what do you do when you do not have the luxury of collecting more data? Then there is something called windowing we will talk about that, what do you do is like you said there is a discontinuity to mitigate the border effects as I said whether signals or countries we have always have border problems. So, the to mitigate the boarder effects what we do is we multiply the given signal with the window function that is gentle at the borders; right now when you are dealing with the finite length signal it is as if there is this infinitely long signal which we have not observed it is a procession and we are watching this procession through a window and after we are done we close the window.

So, it is actually being very brutal at the ends, you right to be more gentle at the ends what people suggested and I do not think there is much left for you in doing research there, there are so much research on this. They suggested that you multiply this signal with a gentle window function and so that it is a taper soft at the borders and there are so many papers that have come out and still people are you know people are deep into it contribute to it. These window functions are named after there are different window functions that you can multiply they named after people who have done excludes all the windows that you seen operating system, apart from that you have all the other windows Blackman, Hanning, Hamming, Kaiser whatever if you come up with the better window you can actually have named after you.

But that is one way of another way of mitigating spectral leakage. So, this concludes the discussion on the deterministic world; tomorrow what we are going to do is we are going to enter the world of random signals and there is only one thing that we will have to do which is learn what is the definition of spectral density in a random signal and then we will come across the familiar Wiener-Khinchin relation. So, I leave you with this list of commands in r that are relevant to what we have discussed until now convolution is obtained by convolve in the base package.

(Refer Slide Time: 24:39)



Conv from the signal package and you want to calculate impulse response given the transfer function operator g of q inverts you can actually specify use impz, again from the signal package freqz f r e q z from the signal package computes FRF and FFT as you already know and then the periodogram. So, we will meet tomorrow.

Thank you.